Deep Learning Macroeconomic Predictions: Multivariate Forecasting Using Long Short-Term Memory (LSTM).

Ryan George Sophia Beronilla

McGill University
{ryan.george2, sophia.beronilla}@mail.mcgill.ca
December 3, 2023

Abstract

This project looks at multivariate forecasting of unemployment rates using Long Short-Term Memory (LSTM) networks, a deep learning framework. Unemployment rates stand as a critical metric, influencing economic health and individual lives. This project employs data sourced from the Federal Reserve Economic Data (FRED), focusing on eight variables, including historical unemployment, real GDP, consumer price index, and others. Comparisons with traditional forecasting models, including Vector Autoregression (VAR) and Moving Average, higlight the LSTM model's ability to 'learn' trends and discard outliers. The findings highlight the potential of deep learning techniques in economic forecasting, while acknowledging existing challenges such as the need for larger datasets.

Contents

1 Introduction

Unemployment rates serve as a critical metric of economic health, exerting profound implications on both individual lives and the overall well-being of societies [13]. Unemployment rate represent the percentage of the labor force actively seeking employment but currently without jobs [13]. As a key macroeconomic indicator, unemployment rates offer insights into the efficiency and robustness of an economy, shaping public policy, business strategies, and societal perceptions [13]. Unemployment as a whole has several adverse effects on the populace, effecting mental health and economic opportunities [12]. Members of the general public, as well as Economists, Policymakers and many others closely monitor these figures as they impact decisions ranging from fiscal or monetary policies to financial investment strategies. Consequently, understanding and accurately forecasting unemployment rates are important for informed decision-making. With the advent of readily available data and machine learning libraries, applications of deep learning method such as Long Short-Term Memory (LSTM) networks in predicting such macroeconomic variables is possible. To this end, some papers have tried implementing LSTM for Unemployment predictions [14], however, a multivariate approach has rarely been tried.

2 Method

2.1 Long Short-Term Memory Networks

Before discussing the implementation of the model for multivariate unemployment prediction, it is necessary to look over the architecture of an LSTM cell. Long Short-Term Memory was developed to combat the vanishing gradient issued faced by Recurrent Neural Networks (RNNs), which can hinder the training of traditional RNNs over long sequences [11]. LSTMs introduce a more complex architecture, consisting of specialized memory cells, to capture and retain information over extended periods [9][11]. The architecture of an LSTM cell is illustrated in Figure 1, in which we see three gates. The first of these gates is the *forget gate* (f_t), which 'decides' what information from the cell state (C_{t-}) should be discarded or kept [11]. It takes as input the previous cell state and the current input, and outputs a value between 0 and 1 for each element in the cell state. Mathematically, this gate can be modelled as:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f$$

Where W_f and b_f are the weight matrix and bias for the forget gate, and σ is the sigmoid activiation function [9]. The previous hidden state is denoted by h_{t-1} while x_t is the current input[11].

The second gate is the *input gate* (i_t). In essence, this gate determines what new information to store in the cell state [9][11]. It consists of two parts, a sigmoid (σ) layer that decides and a tanh layer that creates a vector of new values ($\tilde{C_t}$). This can be seen visually in the middle gate depicted in Figure 1. The input gate is given by:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i$$

While the candidate value is given by

$$\tilde{C_t} = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C$$

W_i, b_i, W_C, b_C are the weights and biases for the gate respectively [11]. Within this cell, the cell state is updated by combining the information from the forget gate and the new candidate values [9][11]. This is given by

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C_t}$$

The last game is the Output Gate (o_t), which determines the next hidden state (h_t) [11]. This will be based on the updated cell states, as the hidden state contains information abotu the current input and past hidden state [11]. The given equations of the output state, and hidden state are as follows:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_0)$$

$$h_t = o_t \cdot \tanh(C_t)$$

Where $W_o, b_o, \tanh$ are the weights (w_o), biases (b_o), and hypoterbolic activation function for the output gate [9][11].

2.2 Model Evaluation Metrics

When applying deep learning models, it is not uncommon to utilize various metrics to evaluate model performance and overall accuracy. In this project, Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE) were used for model evaluation.

Mean Absolute Error evaluates the average difference between the predicted value and the true value. The ideal value of this model is 0, indicating no difference between the predicted value and actual value [14]. The formula for MAE is as follows:

$$MAE = \frac{\sum_{i=1}^{n} |y_i - \hat{y}_i|}{n}$$

Mean Squared Error is a commonly used metric to measure loss, especially with regression applications [14]. MSE can be thought of as "the average frame loss per sample" of the dataset [14].

$$MSE = \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n}$$

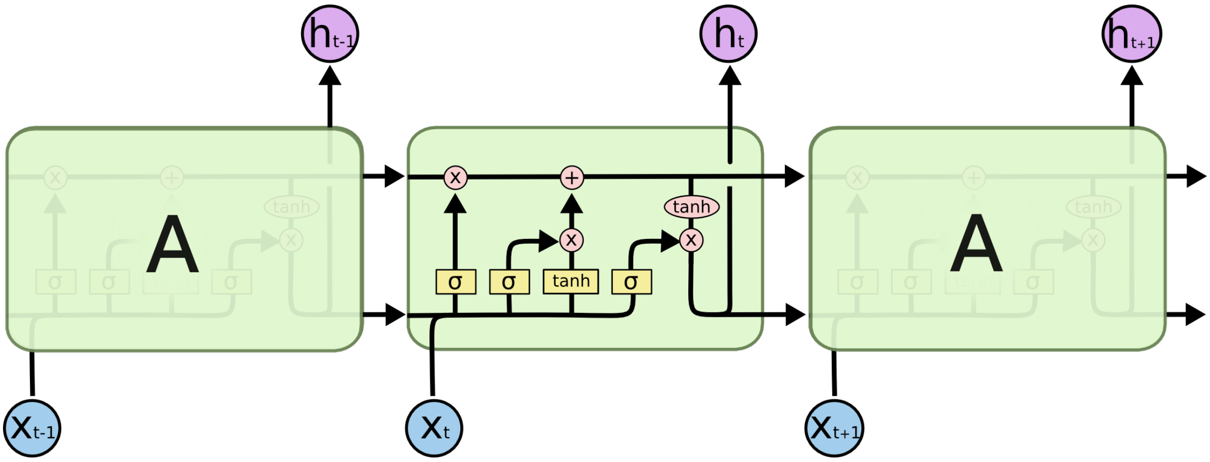

Figure 1: LSTM Architecture, Courtsey of Christopher Olah [11]

Finally, Root Mean Squared Error is simply the square-root of the Mean Squared Error. Taking the square-root of the MSE allows for the units in the RMSE to be the same as the target, which allows for a more intuitive understanding of loss [13]. The equation for RMSE is as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}}$$

2.3 Data

The data used in this model are sourced from the Federal Resource Economic Data (FRED). In particular, eight variables are of interest to us; historical unemployment, real gross domestic product, consumer price index, federal funds effective rates, industrial manufacturing production, labour force participation rate, velocity of M2 money stock, and social security transfers [1–8]. Unemployment rate is a great indicator for the health of an economy, therefore historical unemployment can lend insight for future unemployment trends. Real GDP measures the total value of all goods and services produced by a country within its borders, multiplied by a GDP deflator. Theoretically, as Real GDP increases, unemployment rates should fall as output and participation trend similarly. Consumer Price Index is a basket of goods set out by the central bank, in our case the Federal Reserve. The CPI is a GDP deflator which should have some correlation to Real GDP but not perfectly collinearity. Effective rates represents the interest rates at which banks lend funds to each other overnight. It is a key indicator for short-term interest rates and plays a significant role in monetary policy. Industrial manufacturing production measures the output of goods produced by the manufacturing sector. This is a good measure of economic activity as if output slumps, it is likely that unemployment will follow. Labour force participation rate provides valuable information about the proportion of the population in the labour market. Velocity of M2 money stock represents the speed at which money in an economy exchanges hands. M2 includes currency, as well as deposits in financial intermediaries such as Roth IRAs, as well as money market funds [6]. Finally, social security transfers is a measure of government financial support to individuals such as veterans, retirees and those with disabilities.

It should be noted that the data spans 51 years, segmented by quarters. Specifically, each variable goes from Quarter 1 of 1972 to Quarter 2 of 2023. Figure 2 is a visual representation of each of the variables. We can see from the plots the various economic recessions, the most recent of these shocks is the effects from the 2020 Coronavirus pandemic. Figure 3 is a pearson correlation matrix of our variables, where a measure of 1 indicates a perfect positive correlation and -1 indicates perfect negative correlation. A correlation matrix was used in order to see if there was perfect correlation that could be removed in order to reduce computing time. From the visualization, we can see that there are some variables close to perfect correlation but not quite there. As a result, all of the selected variables will be included to train our LSTM model.

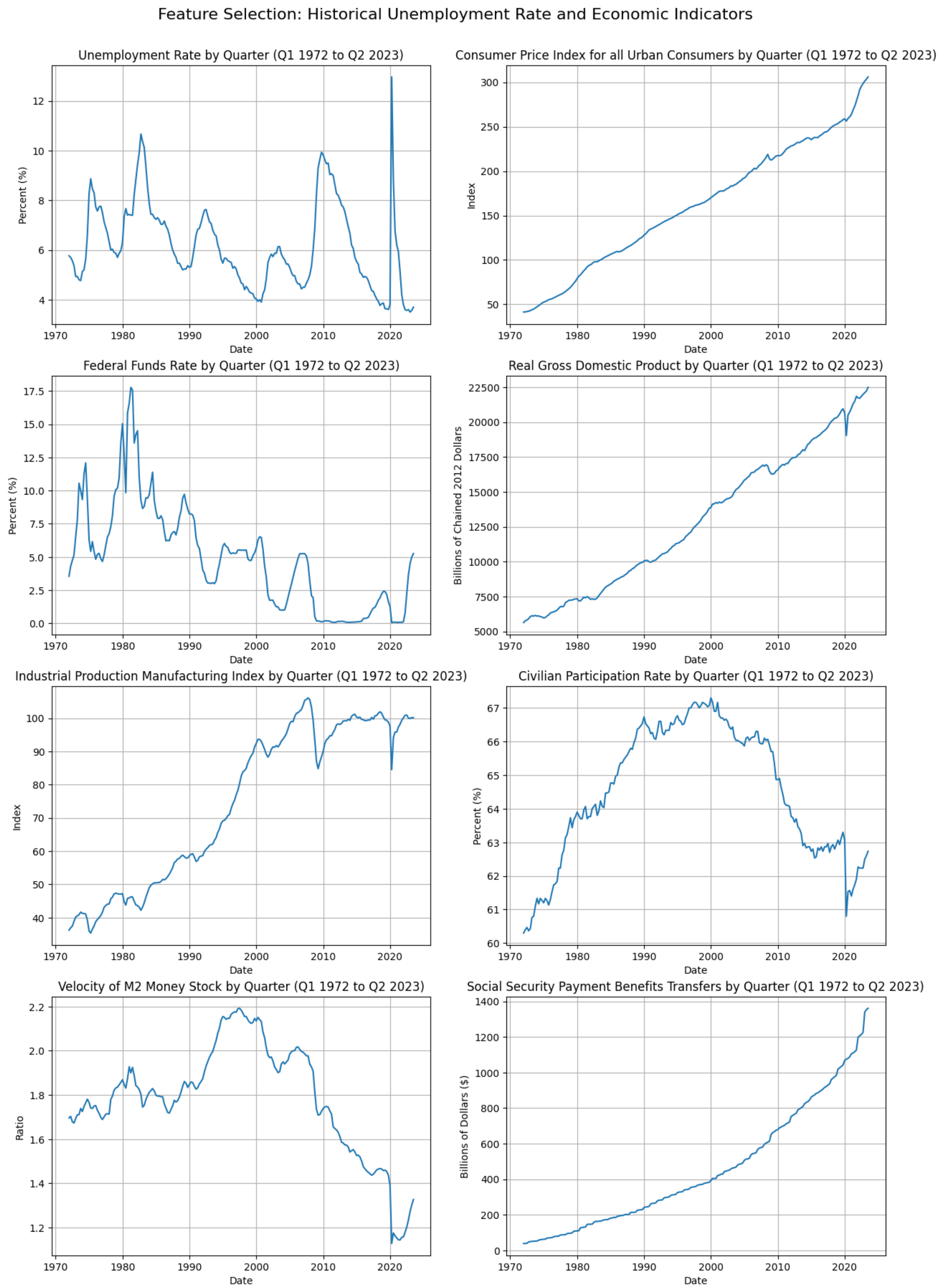

Figure 2: Visualization of Selected Variables (Q1 1972 to Q2 2023).

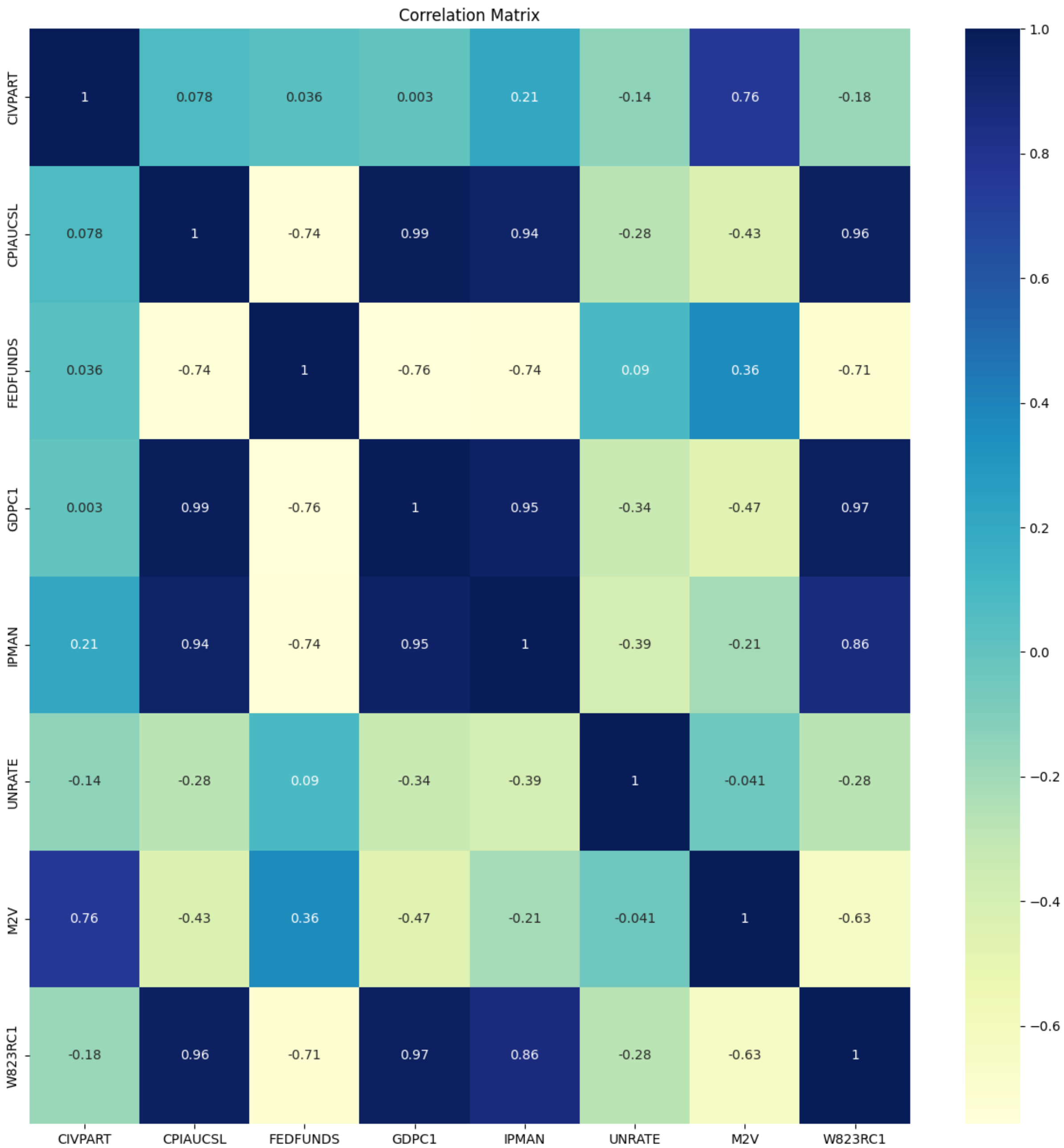

Figure 3: Pearson Correlation Matrix of Selected Variables.

3 Results

3.1 Model Performance

The LSTM model was ran over 30 epochs to avoid over-fitting. Hyperparameter tuning was conducted with the use of a dropout rate of 35%. The model also utilized a robust scaler to minimize the effect of outliers over the 51 years of quarterly data on learning. The activation function was a Rectified Linear Unit (relu), to avoid the potential for vanishing gradients. Figure 4 illustrates the loss for the model. It is obvious from the validation loss that over fitting is a prominent issue with the model. This could be due in part to the lack of data, as there was roughly 200 observations in the entire dataset over 8 variables. Looking at the model performance metrics is fairly promising as the MSE is 0.4248, RMSE is 0.6518, and finally the MAE is 0.5459. When feeding the LSTM model into Keras, we lagged the model quarterly. Therefore, the metrics are really representing an error for 3 months at a time which can compile and cause a higher error rate. With this in mind, the model performed decently well. Figure 6 is a plot of the actual versus predicted unemployment rate, we can see from this visualization that the model is underfitting and overfitting at times. Figure 5 illustrates the forecasted unemployment rate for Quarter 1 2024 to Quarter 4 2024. Based on the predictions versus actual unemployment rate visualization in Figure 6, we expect this forecast to be slightly overpredicted.

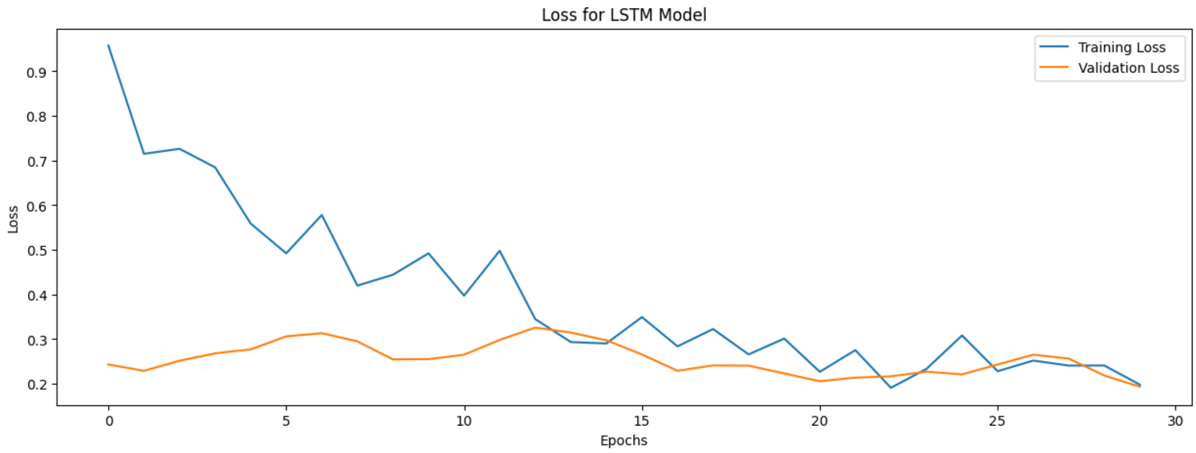

Figure 4: Graph of Loss for LSTM Model Over 30 Epochs.

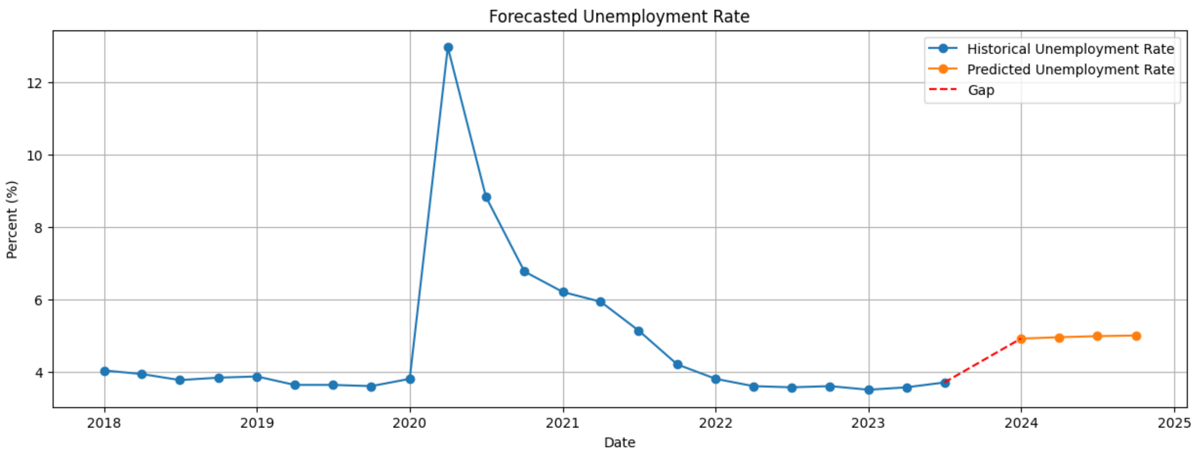

Figure 5: Graph of Forecasted Multivariate Unemployment Rate (Q1 2018 - Q4 2024).

3.2 LSTM in comparison to Moving Average and Vector Autoregression

Although the model has some inaccuracies, it should be compared to commonly used models for macroeconomic forecasting. Vector Autoregression (VAR), is a statistical method used in econometrics and time

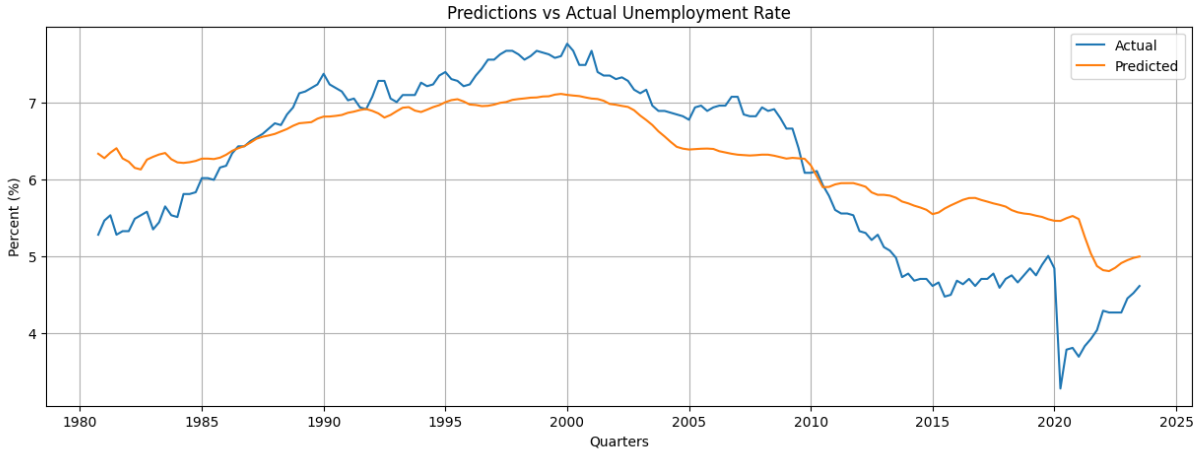

Figure 6: Graph of Predicted versus. Actual Unemployment Rate (1981 - 2023).

series analysis to model the relationships between multiple time series variables [10]. In a VAR model, each variable in the system is modeled as a linear function of its past values and the past values of all other variables in the system [10]. This is represented mathematically as:

$$y_t = c + A_1 y_{t-1} + A_2 y_{t-2} + ... + A_n y_{t-n}$$

Figure 7 plots the predicted values from a VAR to the LSTM predictions. From this graph, we can see that LSTM is consistently higher than the VAR predictions, although the values do come close in Quarter 2 of 2024. It should be noted that given the data and shaping of the LSTM model, it is lagged for an additional quarter that the VAR model is not. The drop-off in the predicted VAR value in Quarter 3 of 2024 does seem unlikely given how "smooth" the historical trends have been from quarter to quarter. Looking at a moving average in comparison to the LSTM shows that the model was able to learn trends and discard outliers, as it follows the actual trend fairly well.

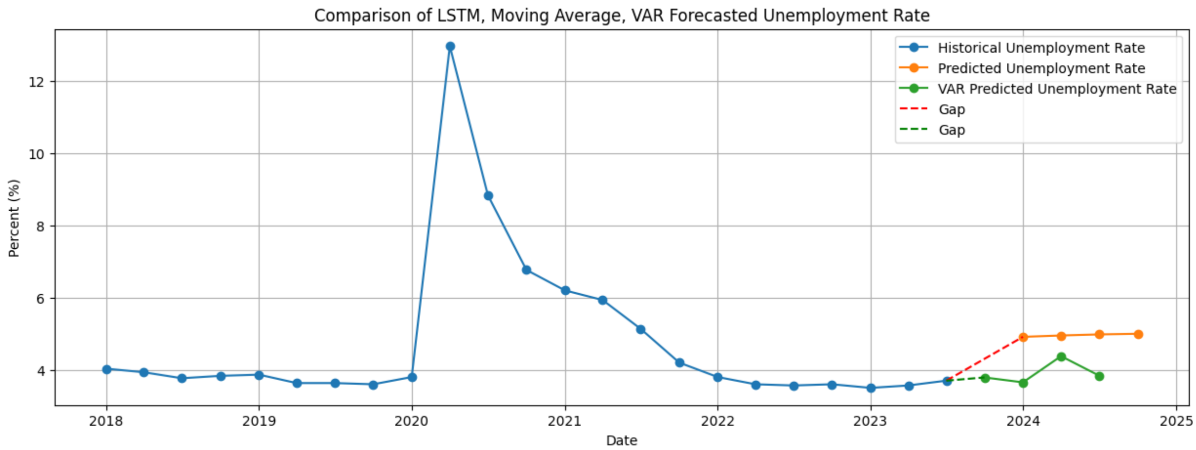

Figure 7: Comparison of LSTM versus. Vector AutoRegression (VAR) Predictions.

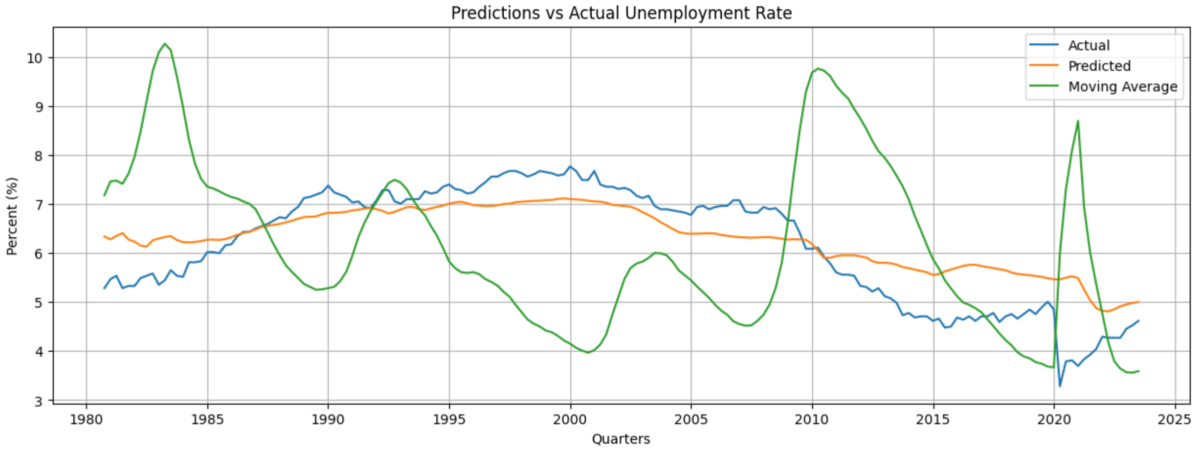

Figure 8: Comparison of Moving Average versus LSTM.

4 Conclusion

In conclusion, the unemployment rate stands as a vital metric in assessing economic health. Its role as a key macroeconomic indicator provides insights into the robustness of an economy, impacting various aspect of society such as public policies to overall societal health. The adverse effects of unemployment on mental health and economic opportunities underscore the significance of accurate predictions for informed decision-making [12]. This study leveraged the power of Long Short-Term Memory (LSTM) networks, a deep learning method, to forecast unemployment rates using multivariate data sourced from the Federal Reserve Economic Data (FRED). While LSTM models have been explored for unemployment predictions, the multivariate approach introduced in this project shows the potential for deep learning in more effective forecasting techniques. Although the model faced overal overfitting issues, there are fixes that can be made to our approach that would potentially provide a more sophisticated method for accurate forecasting. To start, the data being segmented as quarters compiled errors and potentially created false trends that the model may have picked up on. Moving forward, more data is clearly required in order to have more accurate results with higher validity.

References

[1] Labor force participation rate, Nov 2023. URL https://fred.stlouisfed.org/series/CIVPART#0.

[2] Consumer price index for all urban consumers: All items in u.s. city average, Nov 2023. URL https://fred.stlouisfed.org/series/CPIAUCSL#0.

[3] Federal funds effective rates, Nov 2023. URL https://fred.stlouisfed.org/series/FEDFUNDS#0.

[4] Real gross domestic product, Nov 2023. URL https://fred.stlouisfed.org/series/GDPC1#0.

[5] Industrial production: Manufacturing (naics), Nov 2023. URL https://fred.stlouisfed.org/series/IPMAN#0.

[6] Velocity of m2 money stock, Nov 2023. URL https://fred.stlouisfed.org/series/M2V.

[7] Unemployment rate, Nov 2023. URL https://fred.stlouisfed.org/series/UNRATE#0.

[8] Personal current transfer receipts: Government social benefits to persons: Social security, Nov 2023. URL https://fred.stlouisfed.org/series/W823RC1.

[9] Aurélien Géron. *Hands-on machine learning with scikit-learn, Keras, and tensorflow: Concepts, tools, and techniques to build Intelligent Systems*. O'Reilly Media, 2023.

[10] Abdulnasser Hatemi-J. Multivariate tests for autocorrelation in the stable and unstable VAR models. *Economic Modelling*, 21(4):661–683, July 2004. URL https://ideas.repec.org/a/eee/ecmode/v21y2004i4p661-683.html.

[11] Christopher Olah. Understanding lstm networks, Aug 2015. URL https://colah.github.io/posts/2015-08-Understanding-LSTMs/.

[12] Laura Pohlan. Unemployment and social exclusion. *Journal of Economic Behavior & Organization*, 164:273–299, 2019. ISSN 0167-2681. doi:https://doi.org/10.1016/j.jebo.2019.06.006. URL https://www.sciencedirect.com/science/article/pii/S0167268119301969.

[13] Preethi Pratap, Alison Dickson, Marsha Love, Joe Zanoni, Caitlin Donato, Michael A Flynn, and Paul A Schulte. Public health impacts of underemployment and unemployment in the united states: Exploring perceptions, gaps and opportunities, Sep 2021. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8508259/.

[14] Mustafa Yurtsever. Unemployment rate forecasting: Lstm-gru hybrid approach - journal for labour market research, Jun 2023. URL https://labourmarketresearch.springeropen.com/articles/10.1186/s12651-023-00345-8#citeas.

A Source Code.